

Introduction

This project will be perform analysis using the “Arrest and Strip Search” dataset from the Toronto Police Service’s Public Service Data Portal. To exercise adequate research analysis for this dataset, the project will start with a literature review of the background of the data set. This research step is crucial as it identifies the context and current climate of the associated parties, in this instance would be the Toronto Public Service. Afterwhich, this project will construct three (3) research questions to tackle the data from three perspective. To avoid overlapping in the data analytics stage, the three research questions will include difference difference attributes from the datasets.

Afterwhich, to better understand the dataset, the team will conduct exploratory data analysis (EDA) on the dataset to summarize the main characteristics using graphical and other visualization methods. To better understand the parametric results from this analysis, t-test will also be used to compare the means of two groups or samples from the same research question. This is done with the with hypothesis and under the assumption that the data follows a normal distribution and that the variances of the two groups are equal. The t-test will calculates a t-value, which measures the difference between the means of the two groups relative to the variability within each group.

After then, statistical analysis will be conducted using methodologies such as the one-way ANOVA test and two-way ANOVA test. They are used to compare the variation between groups to the variation within groups. ANOVA test will calculate an F-statistic, which is the ratio of the variance between groups to the variance within groups. If the F-statistic is large enough, it indicates that there is a significant difference between the means of the groups. ANOVA can also be used to test for interaction effects between different factors or variables.

Finally, the results of the above tests will be synthesized to draw a conclusion on the hypothesis constructed. After which, to improve our work going forward, a brief reflection will be conducted to summarize areas of improvement for our future works.

Literature Review

Unlawful strip searches have been a long-time concern in public debate in Canada for many years, particularly subjected to minority groups. This literature review will examine the definition of unlawful strip searches, the legal framework governing such searches, and the psychological and emotional impact of such searches on minority groups in Canada.

Unlawful strip searches are identified as searches that involve the removal of some or all of an individual's clothing in order to search for contraband, weapons, or other prohibited items, and are conducted in violation of an individual's rights or without sufficient justification. The Toronto Police further defines strip search as “the removal or rearrangement of some or all of the clothing of a person so as to permit a visual inspection of a person’s private areas, namely genitals, buttocks, breasts ... or undergarments” (Lemke, 2022). This definition is also widely adopted by the Canadian Courts and oversight bodies as a convetional and justified method of searches. A

key characteristic of strip search is that it is conducted in absence “of the consent of the searched person, under the supervision of the police, in police-controlled spaces” (Lemke, 2022).

Officers are generally allowed to conduct strip searches when an individual is being admitted to a correctional facility, but again, the search must be conducted in a manner that is minimally intrusive. However, despite these rules, individuals of marginalized groups in Canada, including Indigenous peoples, Black Canadians, and people with disabilities, have reported experiencing strip searches that are “degrading, humiliating, and unnecessary” (Lemke, 2022). In particular, Indigenous women and girls have been subjected to strip searches in a manner that is discriminatory and harmful, leading to calls for reform of strip search policies in Canada. Furthermore, there are general social concerns over that strip searches are becoming a tool of harassment and discrimination against minority groups. Where police forces are utilizing racial profiling to conduct searches. Researches have found that minority groups such as Black Canadians, Indigenous people, and people with visible disabilities are disproportionately subjected to strip searches when there is no evidence of reasonable suspicions (Lemke, 2022).

A 2019 report conducted by Ontario’s Independent Police Review Director (OIPRD), a review board in which collects and reviews data from police forces across Ontario, revealed “here are approximately 22,000 strip-searches conducted by police each year, with the vast majority taking place in the city of Toronto. More than four out of every 10 people arrested by Toronto police have been ordered to remove their clothing in recent years” (Ghobrial, 2021). Many consider this to be an excess and an alarming amount of strip searches to be conducted on a routinely basis by the Toronto Police. After this revelation, the Toronto Police force quickly went into action and “put in a number of measures, including reviewing our procedures, training and accountability” (Ghobrial, 2021).

What has subsequently revealed is that “the [Toronto Police] force has overhauled when and how officers perform a procedure dubbed ‘inherently degrading’ by Canada’s top court — changes the interim chief says have already led to a significant drop in numbers” (Gillis, 2020). Two main issues which was disclosed that through internal investigation were the outdated training provided to Toronto police and poor data collection process. To combat these issues, there is a new set of mandatory steps to approach strip searches. This means officers cannot perform strip search prior to conducting a less invasive approach, such as frisk search where clothing are not removed from the individual. Furthermore to enhance its data collection process, the Toronto Police Officers now have higher level of requirements for reporting on new strip search. Which includes all legal requirements on how to determine if there are reasonable grounds to authorize a strip search. Once the forms have been complete, they are also be analyzed and reported.

In conclusion, unlawful strip searches are a significant social issue in Toronto and even in Canada, in particular with minority groups who are disproportionately affected. While strip searches may deem to be necessary in certain scenarios, it is important that the police officers are conducting them in a manner that respecting the individual's rights and is minimizing physical

and emotional harms. Further reporting and analysis is needed to better understand the impact of strip searches on minority groups and to form strategic implementations to minimize the negative impacts of such searches on affected individuals, in particularly with the marginalized communities.

Dataset Description

The dataset used to conduct this research project is titled "Arrests and Strip Searches". It is provided by the Toronto Police Service and contains information on arrests and strip searches conducted by the police between 2010 and 2019. The dataset contains 65,276 records and 24 attributes. 12 of which are numerical variables and 12 of which are categorical variables. The numerical variables in the dataset are represented using 1 and 0, where 1 = true or actions have been taken and 0 = untrue or no actions have been taken. The variables in the dataset include the arrest date, time, location, and type of offence. Other variables include the age, gender, and race of the individual arrested, as well as the reason for the arrest and whether a strip search was conducted. The dataset also includes information on the type of strip search conducted, the location of the search, and the outcome of the search.

The dataset can be used to analyze patterns in arrests and strip searches conducted by the Toronto Police Service, including any potential disparities in the use of strip searches against certain demographics. Researchers may also be interested in exploring the reasons for arrest and the outcomes of strip searches, as well as examining any potential biases in the data. We also acknowledge that due to Toronto Police Service's internal documentation issues, there may be records in the data where an individual was strip searched however does not indicate through booking under "booked" (where value = 0). In this scenario, the individual is to be considered that a booking has taken place.

Furthermore, the team is working under the assumption that all records in the provided dataset are considered to be arrested since that all records are assigned a "Arrest_ID". Overall, this dataset provides a valuable resource for researchers, policymakers, and community members who are interested in understanding the use of strip searches by law enforcement in Toronto and its potential impact on vulnerable communities.

Research Question 1

Research Objective & Question

As literature review research supports that, police forces may have practiced discriminatory measures against certain profile of people The first research question will be tackling if there are discriminatory practice conducted in terms of gender.

Research Question 1: Does the Toronto Police Service practice discriminatory measures towards male individuals during strip search?

The data which will be utilized to conduct t-test and ANOVA test will the attributes of "Sex", "StripSearch", and "Arrest_Month". We predict that there is statistically significant relationship between the rate of strip search and male individuals across all time frames. Through this

research questions we also want to unveil whether male individuals are proportionate more likely to be arrest as well.

Data Preprocessing

The relevant data provided by the “Arrests and Strip Searches” by the Toronto Police Service are relatively extensive. To extract the data require related to purpose of this research question. Columns of “Sex”, “Arrest_Month”, “StripSearch”, and “Arrest Year” were selected to create a new dataframe to calculate for the information required. We acknowledge that there are other genders outside of the binary gender assignment. However, for the purpose of this research, all statistical calculations will be base off of Female and Male data available through the Toronto Police Service’s dataset. Therefore, all values in the “Sex” attribute = “U” were not considered as part of T-test and ANOVA test.

EDA

To better approach the data at hand, the first step is to perform descriptive analysis and build visualization to better scope our future research direction. EDAs can help us understand overall trends and provide more context when tackling specific research questions,

Table 1: Female vs. Male Total Strip Search and Total Arrest

Sex	Strip Search	Total Strip Search
Female	0	0
Female	1	1283
Male	0	0
Male	1	6518

The data in table 1 simply lays out the total strip search for female and male individuals according to the dataset provided by the Toronto Police Service. Under the column “Strip Search”, 0 mean there were no strip search conducted and 1 mean there was a strip search conducted.

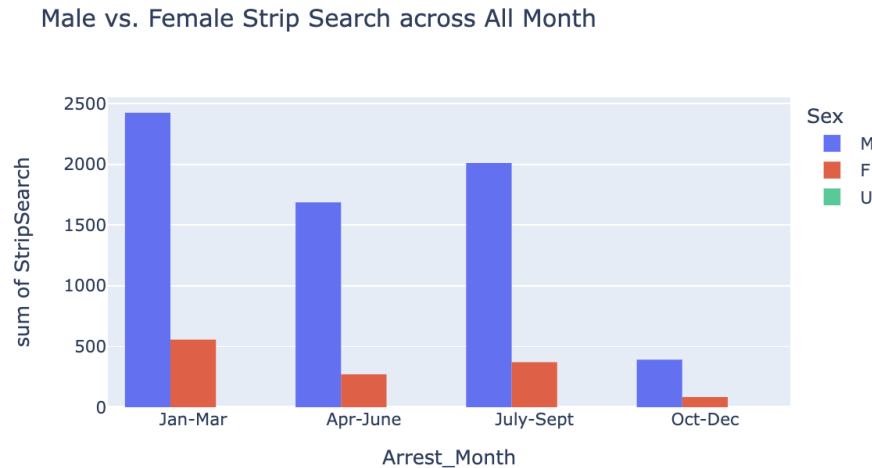
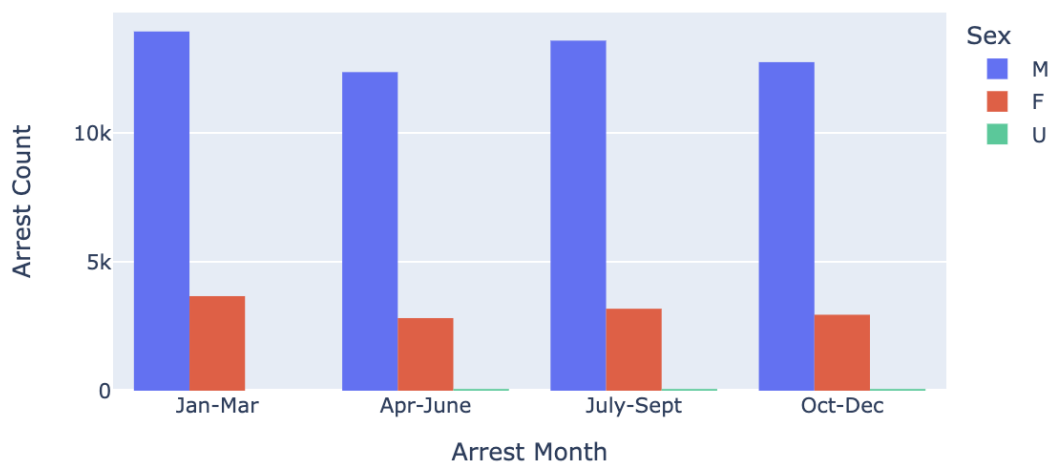


Figure 1: Female Vs. Male Total strip searches across each quarter for 2020 and 2021
To understand the theis significant difference between male and female rate of strip search is a consistent pattern. The above figure uses data from both 2020 and 2021 and visualizes that this difference persists throughout all months of the year. However, it does appear that from October to December, there is a significant drop in both male and female total of strip searches. Although, Sex of “U” were included in the graph, there was no significant data nor relevancy to be considered part of this research question.

Figure 2: Female Vs Male Arrests across each quarter for 2020 and 2021
Female Vs Male Arrest Over Month

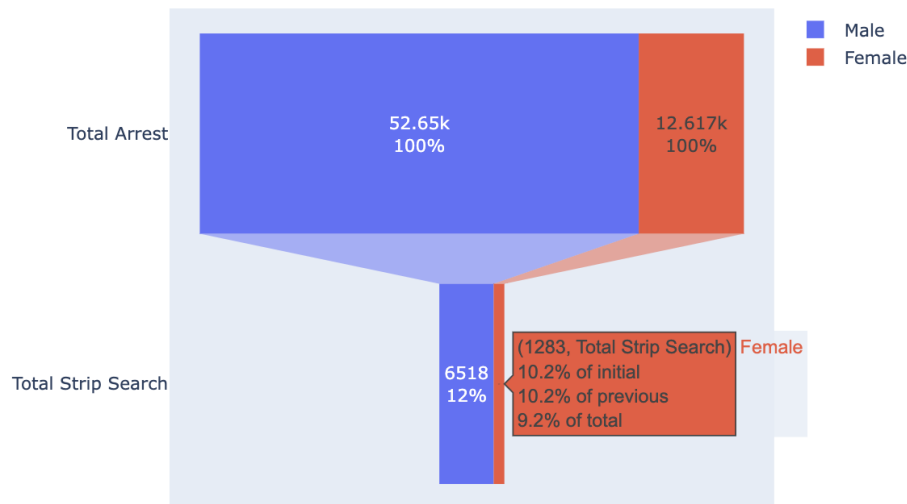


While working under the assumption that all records with an unique “Arrest_ID” will be classified as arrested. It appears that this significant difference between female and male individuals are not only apparent for strip

searches. It is evident in figure 2 that the rate of arrest of male individuals are also significantly higher. This then triggers the question whether does the rate of arrested to the rate of strip search have any significant difference between male and female.

Figure 3: Female vs. Male Total Arrest to Total Strip Search fro 2020 and 2021

Female vs. Male total arrest with Total Strip Search



The funnel in figure 3 shows the rate of percentage of total female vs. male individuals which were arrested and the rate of percentage from that arrested total that performed a strip search. Numerical percentage and total are listed in the below table.

Table 2: Female vs. Male numerical total details on arrested and strip searches and the rate of strip search for 2020 and 2021.

Sex	Strip Search	Total Strip Search	Total Arrested	Rate of Strip Search
Female	0	0	11,334	
Female	1	1283	1,283	10.17%
Male	0	0	46,132	
Male	1	6518	6,514	12.38%

The above results from figure 3 and table 2 shows that although in number the total of male individuals arrested and strip search are significantly higher than female individuals. The percentage of rate of strip search from the arrest total is in comparison have a much less significant difference. The rate of strip search from total arrest for female individuals is 10.17%, where a total of 1283 individuals were strip searched from the total of 12617 individuals that

were arrested. The rate of strip search from total arrest for male individuals are 12.38%, where a total of 6518 individuals were strip searched from the total of 52646 individuals that were arrested.

To collect more statistical evidence and insights into the research questions 1, we will be conducting four t-tests. Each test will measure the mean difference between the rate of strip search between female and male individual from a particular time period within the year. Each t-test will begin with constructing a hypothesis and is accompanied by assumptions .

T-test

Table 3: t-test result (rate of strip search for female and male individual from January - March)

t-test	
t-statistic	p-value
-3.19435268633	0.00140391577

Null hypothesis (H_0): There are NO significant difference between the rate of strip search for female and male individuals from January - March. Alternative hypothesis (H_1): There are significant difference between the rate of strip search for female and male individuals from January - March

The results of the t-test implies that to reject null hypothesis. The p-value is less than 0.05, which indicates the mean of female and male individuals that have been strip search from January to March for 2020 and 2021 are significant different. The result of t-statistic for this hypothesis is -3.19, this indicates the mean of female individuals that have been strip searched is less than the mean of male individuals that have been strip searched from January to March. The results also indicates that the difference between the means between the two groups is quite large, relative to the variability within each group. To further interpret the result, this mean of female individuals are more than three standard errors below the mean of male individuals, making them significantly different from each other.

Table 4: t-test result (rate of strip search for female and male individual from April - June)

t-test	
t-statistic	p-value
-5.764445755134285	8.351787360000213e-09

Null hypothesis (H_0): There are NO significant difference between the rate of strip search for female and male individuals from April - June.

Alternative hypothesis (H_1): There are significant difference between the rate of strip search for female and male individuals from April - June

The results of the t-test implies that to reject null hypothesis. The p-value is less than 0.05, which indicates the mean of female and male individuals that have been strip search from April - June for 2020 and 2021 are significant different. The result of t-statistic for this hypothesis is -5.76, this indicates the mean of female individuals that have been strip searched is less than the mean of male individuals by more than five standard errors.

Table 5: t-test result (rate of strip search for female and male individual from July - September)

t-test	
t-statistic	p-value
-4.5804355197644275	4.6734838605815446e-06

Null hypothesis (H_0): There are NO significant difference between the rate of strip search for female and male individuals from July - September.

Alternative hypothesis (H_1): There are significant difference between the rate of strip search for female and male individuals from July - September

The results of the t-test implies that to reject null hypothesis. The p-value is less than 0.05, which means the means of female and male individuals that have been strip search from July - September for 2020 and 2021 are significant different. The result of t-statistic for this hypothesis is -4.58, this indicates the mean of female individuals that have been strip searched is less than the mean of male individuals by more than four standard errors.

Table 6: t-test result (rate of strip search for female and male individual from Oct - Dec)

t-test	
t-statistic	p-value
-0.6403933061883674	0.5219262800514645

Null hypothesis (H_0): There are NO significant difference between the rate of strip search for female and male individuals from October - December.

Alternative hypothesis (H_1): There are significant difference between the rate of strip search for female and male individuals from October - December.

Fail to reject null hypothesis: There is not enough evidence to conclude that the means of the two groups are different. The result of t-statistic for this hypothesis is -0.64, this indicates the mean of female individuals that have been strip searched is less than the mean of male individuals by 0.64.

Research Method

A one-way ANOVA test was conducted for the relationships between rate of strip search and gender. “Sex” is the independent variable, and “StripSearch” is the dependent variable.

Null hypothesis (H_0): The rate of strip search is NOT different between female and male.

Alternative hypothesis (H_1): The rate of strip search is different between female and male.

Table 7: One-way ANOVA test (Sex & Strip Search)

	sum_sq	mean_sq	F	PR(>F)
Sex	1.434810	1.434810	10.203889	0.001404
Residual	2476.353755	0.140614	NaN	NaN

The results from the ANOVA test for the F-statistic is 10.203889, indicates the variation explained by “Sex” is significantly larger than the variation that is not explained by “Sex”. The PR(>F) shows the p-value for the F-statistic. For “Sex”, the PR(>F) is 0.001404, which is the probability of observing F-statistics as large as 10.203889. Overall, there is evidence that the means of the dependent variable (strip search) are different for female and male.

Table 8: Tukey’s HSD test result (Strip search for Female vs. Male)

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
Group 1	Group 2	meandiff	p-adj	lower	upper	reject
Female	Male	0.0222	0.0014	0.0086	0.0359	TRUE

The results from the table 8 indicate the mean difference between females and males rate of strip search is 0.0222 and the adjusted p-value for this comparison is 0.0014, which is below the significance level of 0.05. This suggests that there is a statistical significant difference between the two groups. The confidence interval for the mean difference ranges from 0.0086 to 0.0359. Therefore, the null hypothesis of no difference is rejected. It further suggests that the mean for rate of strip search for male is higher than female individuals. Which is consistent with the previous t-test and ANOVA tests.

Results

From the EDAs and descriptive analysis alone, it proves that total arrested and strip searched male individuals are significantly higher than the total arrested and strip searched female individuals. However, the rate of strip search from total arrest does not show as of a significant difference. Accordingly to the table 2 above, there is approximately 2.21% higher chance of male individuals being stripsearched in comparison to female individuals.

The T-test results and ANOVA test results show that the mean the rate of strip search for male individuals are higher than the rate of strip search for female for all periods in a year except for October - December. The overall the results from the tests tells us that there is a higher chance of male individuals being strip search after being arrest than a female individual. However, this is

not substantial evident to conclude that the Toronto Police Service is practicing discriminatory measures toward male individuals. However, It is a question to be asked and conduct further investigations.

Research Question 2

Research Objective & Question

To investigate whether the Toronto Police Service discriminates against people of different ages when making arrest decisions, we decided to conduct t-tests and one-way ANOVA tests on the variables, 'Booked', 'Age_group__at_arrest_', and 'Youth_at_arrest__under_18_years'. We predicted statistically significant relationships between arrest rates and age groups, especially middle-aged people with youth and the elderly. We want to see if the judges were stricter with middle-aged adults and more lenient with teenagers and the elderly because their age was considered to make a judgment. Therefore, our research question is 'Does the Toronto Police Service practice discriminatory measurement towards youth and elderly ('Age_group__at_arrest_', 'Youth_at_arrest__under_18_years') when making arrest decisions ('Booked')?'

Data Preprocessing

Since 'Age_group__at_arrest_' and 'Youth_at_arrest__under_18_years' have duplicated categories, we have to first preprocess the data.

For 'Age_group__at_arrest_', 'Aged 17 years and under', 'Aged 17 years and younger' and 'Aged 65 and older', 'Aged 65 years and older' have the same meaning, so we merge them into the same group respectively. Besides, there were 24 blank records, which is a small number compared to a total of 65,276 records, so we decided to delete them. A new column, 'Age', is created with the cleaned data of 'Age_group__at_arrest_'.

For 'Youth_at_arrest__under_18_years', 'Youth (aged 17 and younger)' and 'Youth (aged 17 years and under)' share the same meaning, so we also integrate them into one group. A new column, 'Youth', is created with the cleaned data of 'Youth_at_arrest__under_18_years'.

EDA

To get a basic understanding of these three variables, we performed a detailed descriptive analysis.

Table 9: Frequency of youth and arrest decisions

Booked/Youth	Not a youth	Youth	Total
0	29,533	1,814	31,347
1	32,677	1,228	33,905

Total	62,210	3,042	65,252
-------	--------	-------	--------

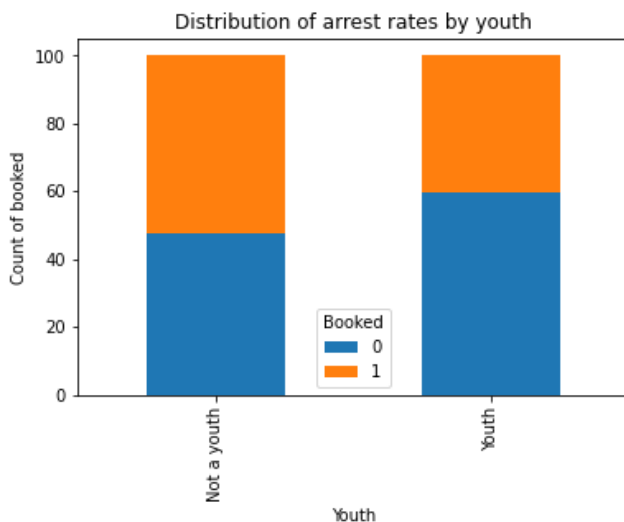
In the dataset, more than 95% were adults, while only a few were youth, and the total number of arrests (1) was slightly higher than the number of non-arrests (0).

Table 10: Percentage of arrest rates grouped by youth

Youth/Booked	0	1
Not a youth	47.473075	52.526925
Youth	59.631821	40.368179

Since the number of youths was too small to distinguish between arrest rates of youth and non-youth, we presented percentage statistics for ‘Youth’ and ‘Not a youth’. Table 10 showed that more adults were booked for crimes, while it was vice versa for the youth.

Figure 4: Distribution of arrest rates by youth



We visualized table 10 as a bar chart to better observe the differences in arrest rates for youth and adults. Figure 4 showed that the arrest rates of adults were relatively higher than that of the youth.

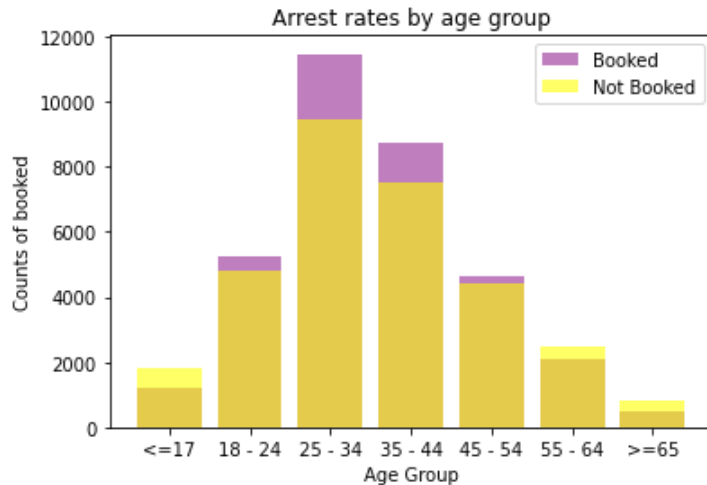
Table 11: Frequency of age groups and arrest rates

	17 and younger	18 - 24	25 - 34	35 - 44	45 - 54	55 - 64	65 and older	Total
0	1,814	4,807	9,480	7,512	4,422	2,485	827	31,347
1	1,228	5,234	11,469	8,730	4,644	2,105	495	33,905

Total	3,042	10,041	20,949	16,242	9,066	4,590	1,322	65,252
-------	-------	--------	--------	--------	-------	-------	-------	--------

In the dataset, most records were from middle-aged people aged 25 and 44, while the fewest were from youth (17 and younger) and elderly (65 and older).

Figure 5: Bar chart of arrest rates by age group



We visualized table 11 as a bar chart to better understand the distribution of arrested and non-arrested people. Since people aged between 25 and 34 and 35 and 44 were the most populous in the dataset, they made up the majority of both booked or not booked cases. It also showed that middle-aged people (i.e., 18 – 24, 25 – 34, 34 – 44, 45 – 54) were more likely to be arrested. The largest disparity in arrests for a crime was among those aged 25 to 34, followed by those aged between 35 and 44. Meanwhile, teenagers (i.e., <=17) and the elderly (i.e., >=55) showed opposite results. Hence, these tables and figures align with our hypothesis that judges may consider age when making arrest decisions, especially for youth and the elderly.

To gain more insight into these three variables, we conducted three t-tests. These tests are consistent with the t-test assumptions as they are independent groups and we want to know their differences in arrest rates.

Table 12: t-test result (Arrest rates of youth and non-youth)

t-test	
t-statistic	p-value
-13.33250828223915	1.492891175480631e-39

Null hypothesis (H_0): There is no significant difference in arrest rates between youth and adults.
Alternative hypothesis (H_1): There is a significant difference in arrest rates between youth and adults.

The p-value is extremely small, $1.493e-39$, implying a statistically significant difference in the average arrest rates between youth and non-youth. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the arrest rates were different between youth and non-youth. Also, the t-statistic is -13.33 , which means that youth had a lower average arrest rate than non-youth. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that judges treat youth and non-youth differently when making arrest decisions.

Table 13: t-test result (Arrest rates of people aged 25 – 34 and 35 – 44)

t-test	
t-statistic	p-value
1.9153617697970984	0.05545454952532147

Null hypothesis (H_0): There is no significant difference in arrest rates between people aged 25 – 34 and 35 – 44.

Alternative hypothesis (H_1): There is a significant difference in arrest rates between people aged 25 – 34 and 35 – 44.

The p-value is 0.055, which is slightly larger than the significance level of 0.05. There does not have a statistically significant difference in the average arrest rates between people aged between 25 – 34 and 35 – 44, which means the null hypothesis cannot be rejected. We do not have significant evidence to conclude that the arrest rates were different between the two middle-aged groups. The small t-statistic also shows similar differences in arrest rates between the two groups. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that judges treated both age groups equally when making arrest decisions.

Table 14: t-test result (Arrest rates of people aged 25 – 34 and 65 and older)

t-test	
t-statistic	p-value
12.582080069704146	$1.3732423394937862e-34$

Null hypothesis (H_0): There is no significant difference in arrest rates between middle-aged people and the elderly.

Alternative hypothesis (H_1): There is a significant difference in arrest rates between middle-aged people and the elderly.

The p-value is extremely small, $1.373e-39$, implying a statistically significant difference in the average arrest rates between people aged 25 – 34 and 65 and older. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the arrest rates were different between

middle-aged people and the elderly. Also, the t-statistics is 12.582, which means that people aged 25 – 34 had a higher average arrest rate than the elderly. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that judges were more considerate of the elderly when making arrest decisions.

Therefore, from these three t-tests, we know that the average arrest rates for youth and non-youth, and middle-aged people and the elderly are statistically significantly different. Yet, the two middle-aged groups exhibit similar mean arrest rates.

Research Method

A one-way ANOVA test was conducted to investigate relationships and differences between arrest rates and age groups. ‘Age’ is the independent variable, and ‘Booked’ is the dependent variable.

Null hypothesis (H_0): The arrest rate does not differ between different age groups.

Alternative hypothesis (H_1): The arrest rate differs between different age groups.

Table 15: One-way ANOVA test result (Age and Booked)

	sum_sq	df	F	PR(>F)
Age	107.805412	6.0	72.452717	1.937452e-90
Residual	16180.124996	65245.0	NaN	NaN

The one-way ANOVA test compared the differences in average arrest rate between 7 independent age groups to see if they were statistically significant. The F-statistic shows a large effect size of 72.452. The corresponding p-value is extremely low at 1.93e-90, which indicates a significant effect of different age groups on arrest rates. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the arrest rates differed significantly between at least 2 age groups. It also showed a strong relationship between age groups and arrest rates.

Since the mean arrest rates were statistically significantly different between age groups, we performed Tukey’s HSD test to measure the differences between each two age groups.

Table 16: Tukey’s HSD test result (Age and Booked)

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	p-adj	lower	upper	reject
17 and younger	18 - 24	0.1176	0.001	0.0872	0.148	True
17 and younger	25 - 34	0.1438	0.001	0.1153	0.1723	True
17 and younger	35 - 44	0.1338	0.001	0.1048	0.1628	True
17 and younger	45 - 54	0.1086	0.001	0.0778	0.1393	True
17 and younger	55 - 64	0.0549	0.001	0.0206	0.0893	True
17 and younger	65 and older	-0.0292	0.5511	-0.0776	0.0191	False
18 - 24	25 - 34	0.0262	0.001	0.0084	0.044	True
18 - 24	35 - 44	0.0162	0.1356	-0.0024	0.0349	False
18 - 24	45 - 54	-0.009	0.8644	-0.0303	0.0123	False
18 - 24	55 - 64	-0.0627	0.001	-0.0888	-0.0365	True
18 - 24	65 and older	-0.1468	0.001	-0.1898	-0.1039	True
25 - 34	35 - 44	-0.01	0.4708	-0.0253	0.0054	False
25 - 34	45 - 54	-0.0352	0.001	-0.0537	-0.0168	True
25 - 34	55 - 64	-0.0889	0.001	-0.1128	-0.0649	True
25 - 34	65 and older	-0.173	0.001	-0.2147	-0.1314	True
35 - 44	45 - 54	-0.0253	0.0021	-0.0445	-0.006	True
35 - 44	55 - 64	-0.0789	0.001	-0.1034	-0.0543	True
35 - 44	65 and older	-0.1631	0.001	-0.2051	-0.1211	True
45 - 54	55 - 64	-0.0536	0.001	-0.0802	-0.027	True
45 - 54	65 and older	-0.1378	0.001	-0.181	-0.0946	True
55 - 64	65 and older	-0.0842	0.001	-0.13	-0.0383	True

Tukey's HSD test showed significant differences in arrest rates across age groups. People aged 17 and younger were significantly different from those aged 18 to 64, but not those aged 65 and older. It may imply that judges treated youth, adults, and the elderly differently when making arrest decisions. Of all statistically significant results, people aged 17 and younger and 65 and older had the greatest differences in average arrest rates from other age groups, with a mean greater than 0.1. Although many other groups achieved statistically significant results, their average arrest rates differed by less than 0.1.

Result

From the descriptive analysis, t-tests, one-way ANOVA test, and Tukey's HSD test, we can know that the mean arrest rates for youth (i.e., 17 and younger) and the elderly (i.e., 65 and older) were significantly different from those for middle-aged people. Meanwhile, there was no evidence of a difference in arrest rates between youth and the elderly. Therefore, these results help answer our research question, 'Does the Toronto Police Service practice discriminatory measurement towards youth and elderly when making arrest decisions?' by telling us that a different relationship between youth and the elderly in the decision to be classified as an offender compared to other middle-aged groups.

However, these analyses did not demonstrate a causal relationship between youth and the elderly with arrest rates. This relationship might also be caused by other factors, such as the type of crime. It is possible that youth and the elderly committed less serious crimes than middle-aged people, so they were less likely to be booked for crimes. Besides, judges might be more lenient with youth and the elderly simply because of their age, so they would have lower arrest rates. In these scenarios, we can observe the same results, with statistically significant differences in arrest rates between youth and middle-aged people and between elderly and middle-aged people, in which youth and the elderly had lower average arrest rates.

Research Question 3

Research Objective & Question

To investigate whether the Toronto Police Service discriminates against minority groups, including youth, the elderly, and people of color, when conducting strip searches, we performed t-tests and two-way ANOVA test on the variables, 'StripSearch', 'Age_group__at_arrest_', and 'Perceived_Race'. We predicted a statistically significant relationship between age group, race, and strip searches. The average number of strip searches may be higher for teenagers and people of color, while the elderly may have significantly fewer strip searches, who are the usual targets of public stereotypes. We want to see if the police were more discriminatory or selective toward these minority groups. Therefore, our research question is 'Does the Toronto Police Service practice discriminatory measurement) towards minority groups ('Age_group__at_arrest_', 'Perceived_Race') when conducting strip searches?'

Data Preprocessing

Since 'Age_group__at_arrest_' was already cleaned up in research question 2, we would use the cleaned data 'Age' in this research question. Besides, 'StripSearch' is a clean column, so we do not have to create another new column.

There are 4 blank records in the 'Perceived_Race' column, which is an extremely small number compared to the entire dataset. Hence, we decided to remove these records.

Also, a new column, 'POC' has been created to better understand the discriminatory outcomes for people of color. It divides race into people of color and white people.

EDA

To gain more insights of these three variables, we conducted a descriptive analysis.

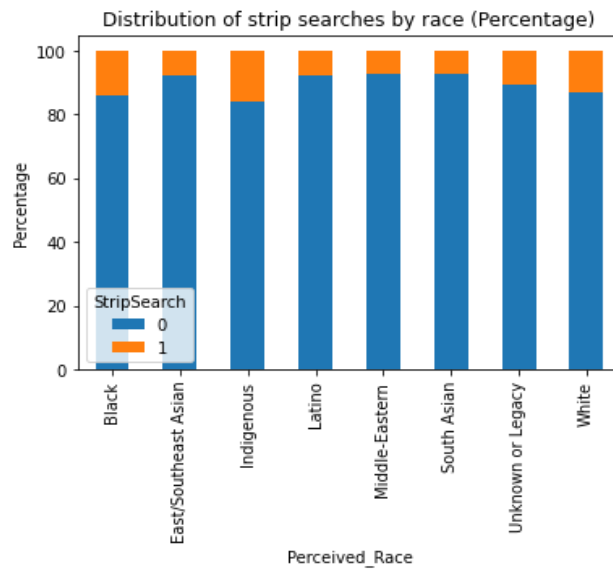
Table 17: Frequency of race groups and strip searches

	0	1	Total
Black	15,084	2,434	17,518
East/Southeast Asian	4,071	341	4,412
Indigenous	1,626	306	1,932
Latino	1,636	132	1,768
Middle-Eastern	3,009	228	3,237
South Asian	3,356	257	3,613
Unknown or Legacy	4,519	536	5,055
White	24,147	3,566	27,713

Total	57,448	7,800	65,248
-------	--------	-------	--------

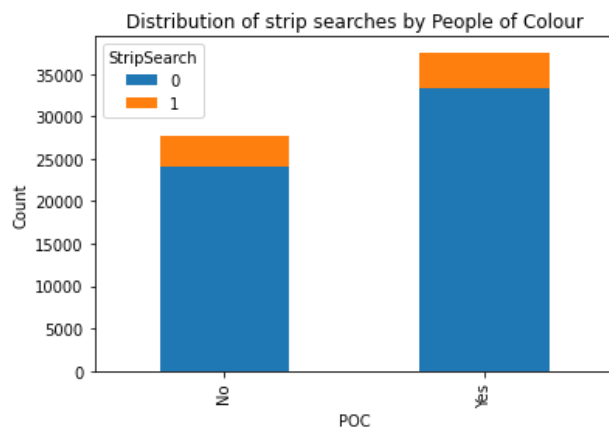
About 90% of the people in the dataset had not experienced a strip search, and only 10% had. Meanwhile, there were significant differences in ethnic distribution. White people contributed more than 40% of the dataset, and the remaining 7 racial groups made up the remaining 60%. The Black and Whites had the most strip searches. Hence, we need to revise this table to percentages to better visualize the differences in strip searches across races.

Figure 6: Distribution of strip searches by race (Percentage)



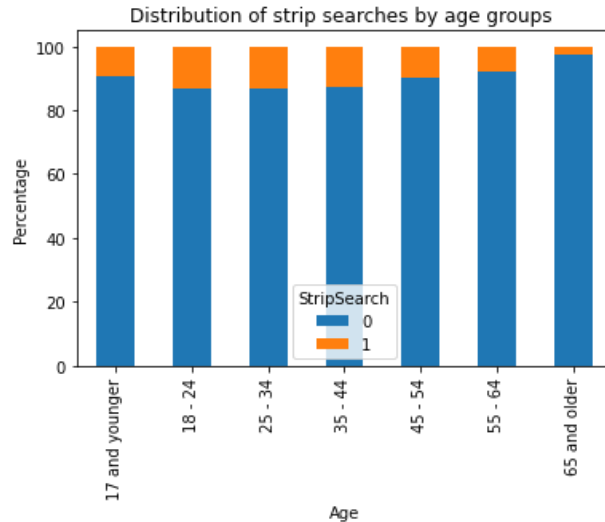
We performed percentage statistics to normalize each ethnic group to see the percentage distribution of their strip searches. Figure 6 showed that all groups experienced similar strip search percentages, with Indigenous, Black, and White experiencing the highest rates. There was no apparent discrimination against specific racial groups when conducting strip searches.

Figure 7: Distribution of strip searches by People of Colour



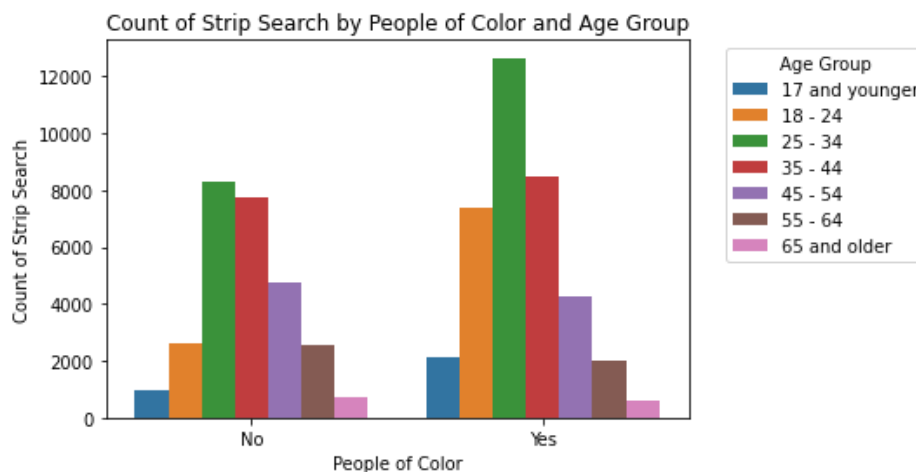
Despite there being more people of color in the dataset, Figure 7 shows that the number of strip searches was similar between the two groups.

Figure 8: Distribution of strip searches by age groups (Percentage)



As with the ethnic groups, the frequency of each age group was uneven, with people aged 18 to 44 contributing more than 70% of the dataset. Hence, we have to translate them into percentages to better understand the distribution of strip searches per age group. The percentage statistics show that people aged 18 to 44 had the highest strip searches while youth (i.e., 17 and younger) and the elderly (i.e., 65 and older) had the lowest. It partly aligns with our assumption that the elderly had the fewest strip searches but is inconsistent with the hypothesis that youth had significantly more strip searches.

Figure 9: Strip searches by people of color and age group



We visualized race, age, and strip searches as a bar chart to investigate their relationship. Yet, due to the uneven distribution of race, we had to divide it into people of color and whites to observe how their strip searches differed across age groups. We can observe that people of color

had higher strip searches for younger people aged 44 and under while the White had slightly higher strip searches for people aged 45 and above. Of all age groups, those between 25 and 34 had the largest disparity in strip searches for people of color and whites. Hence, it does not suggest that strip searches were significantly higher among people of color than whites.

To gain further insight into these variables, we performed three t-tests to measure their differences in strip searches. 'StripSearch' is the dependent variable, while 'Age' and 'Race' are the independent variables. The two independent variables are in independent groups, and we assumed that they did not have equal variances.

Table 18: t-test result (Strip searches for people of color and White people)

t-test	
t-statistic	p-value
-6.1273703503908274	8.991881959699512e-10

Null hypothesis (H_0): There is no significant difference in the number of strip searches between people of color and White people.

Alternative hypothesis (H_1): There is a significant difference in the number of strip searches between people of color and White people.

The p-value is extremely small, 8.992e-10, implying a statistically significant difference in the average number of strip searches between people of color and White people. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the number of strip searches was different between people of color and White people. Also, the t-statistics is -6.127, which means that people of color had lower mean strip searches than White people. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that police treat people of color and White people differently when conducting strip searches, with more strip searches for White people compared to other people of color.

Table 19: t-test result (Strip searches for youth and non-youth)

t-test	
t-statistic	p-value
-5.338693868278733	9.970229086491259e-08

Null hypothesis (H_0): There is no significant difference in the number of strip searches between youth and non-youth.

Alternative hypothesis (H_1): There is a significant difference in the number of strip searches between youth and non-youth.

The p-value is extremely small, $9.97e-10$, implying a statistically significant difference in the average number of strip searches between youth and non-youth. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the number of strip searches was different between youth and non-youth. Also, the t-statistics is -5.339 , which means that youth had lower mean strip searches than non-youth. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that police treat youth better than non-youth when strip searches were considered.

Table 20: t-test result (Strip searches for the elderly and non-elderly)

t-test	
t-statistic	p-value
-20.216171577489664	$7.582399768883804e-81$

Null hypothesis (H_0): There is no significant difference in the number of strip searches between the elderly and non-elderly.

Alternative hypothesis (H_1): There is a significant difference in the number of strip searches between the elderly and non-elderly.

The p-value is extremely small, $7.58e-81$, implying a statistically significant difference in the average number of strip searches between the elderly and non-elderly. Hence, the null hypothesis is rejected. We have significant evidence to conclude that the number of strip searches was different between the elderly and non-elderly. Also, the t-statistics is -20.216 , which means that the elderly had significantly lower mean strip searches than the non-elderly. While there could be many other factors, such as the type of crime, contributed to this result, it might imply that police treat the elderly better than the non-elderly when strip searches were considered.

Therefore, from the three t-tests, we discovered that the minority groups, including the elderly, youth, and people of color, encountered significantly fewer strip searches, which is inconsistent with our prediction that the Toronto Police Service might discriminate against these groups when considering strip searches.

Research Method

A two-way ANOVA test was conducted to investigate the relationships between strip searches, race, and age groups. 'StripSearch' is the dependent variable, while 'Age' and 'Race' are the independent variables.

Null hypotheses (H_0):

1. The strip search does not differ between different age groups.
2. The strip search does not differ between different races.
3. There is no interaction between age and race groups.

Alternative hypotheses (H_1):

1. The strip search differs between different age groups.
2. The strip search differs between different races.
3. There is an interaction between age and race groups.

Table 21: Two-way ANOVA test result (Strip Search, Age, Race)

	sum_sq	df	F	PR(>F)
Perceived_Race	39.840514	7.0	54.690597	1.970197e-78
Age	31.159430	6.0	49.902698	1.523262e-61
Perceived_Race:Age	11.589455	42.0	2.651548	3.477303e-08
Residual	6784.354527	65192.0	NaN	NaN

The two-way ANOVA test compared the differences in mean strip searches between 8 races and 7 age groups to see if they were statistically significant. The F-statistics for the races and age groups were as high as 54.69 and 49.9, respectively, meaning large differences within each group. Besides, the p-values of both races and age groups are less than the significance level of 0.05, so null hypotheses 1 and 2 are rejected. We have significant evidence to conclude that the strip searches differed significantly between at least 2 age and race groups respectively. The interaction between races and age also results in a low p-value, which rejects the third null hypothesis. Besides, the residual sum of squares is 6784, which is relatively large. It indicates that there are many unexplained noises in the model. Therefore, the test result suggests that races, age groups, and their interactions are significant predictors of strip searches. However, more variables are needed to explain the noises, such as the type of crime.

Since mean strip searches were statistically significantly different between age and race groups, we performed two Tukey's HSD tests to observe the differences in the strip searches between each of these two groups.

Table 22: Tukey's HSD test result (Race and Strip Search)

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-0.0617	0.001	-0.0782	-0.0451	True
Black	Indigenous	0.0194	0.1921	-0.0041	0.0429	False
Black	Latino	-0.0643	0.001	-0.0887	-0.0398	True
Black	Middle-Eastern	-0.0685	0.001	-0.0873	-0.0497	True
Black	South Asian	-0.0678	0.001	-0.0857	-0.0499	True
Black	Unknown or Legacy	-0.0329	0.001	-0.0486	-0.0173	True
Black	White	-0.0103	0.0226	-0.0197	-0.0008	True
East/Southeast Asian	Indigenous	0.0811	0.001	0.0543	0.1078	True
East/Southeast Asian	Latino	-0.0026	0.9	-0.0302	0.025	False
East/Southeast Asian	Middle-Eastern	-0.0069	0.9	-0.0295	0.0158	False
East/Southeast Asian	South Asian	-0.0062	0.9	-0.0282	0.0158	False
East/Southeast Asian	Unknown or Legacy	0.0287	0.001	0.0085	0.0489	True
East/Southeast Asian	White	0.0514	0.001	0.0355	0.0673	True
Indigenous	Latino	-0.0837	0.001	-0.116	-0.0515	True
Indigenous	Middle-Eastern	-0.0879	0.001	-0.1161	-0.0598	True
Indigenous	South Asian	-0.0873	0.001	-0.1149	-0.0596	True
Indigenous	Unknown or Legacy	-0.0524	0.001	-0.0786	-0.0261	True
Indigenous	White	-0.0297	0.0024	-0.0528	-0.0066	True
Latino	Middle-Eastern	-0.0042	0.9	-0.0332	0.0248	False
Latino	South Asian	-0.0035	0.9	-0.032	0.0249	False
Latino	Unknown or Legacy	0.0314	0.0106	0.0043	0.0585	True
Latino	White	0.054	0.001	0.03	0.0781	True
Middle-Eastern	South Asian	0.0007	0.9	-0.023	0.0244	False
Middle-Eastern	Unknown or Legacy	0.0356	0.001	0.0135	0.0577	True
Middle-Eastern	White	0.0582	0.001	0.04	0.0765	True
South Asian	Unknown or Legacy	0.0349	0.001	0.0135	0.0563	True
South Asian	White	0.0575	0.001	0.0402	0.0749	True
Unknown or Legacy	White	0.0226	0.001	0.0076	0.0376	True

Table 22 shows the differences in strip searches across race groups. Most Asian groups were not statistically significantly different from other racial groups in strip searches. Although there were many groups whose results were statistically significantly different, their mean differences were similar and within 0.1. Strip searches were statistically significantly different in all 6 comparisons between White people and people of color. Among all of them, the number of strip searches was significantly higher for White people compared to East/Southeast Asians, Latinos, Middle Easterners, and South Asians. For Black and Indigenous people, the results were reversed. Although we cannot conclude that there was a causal relationship between race and strip searches, most people of color were not subjected to more strip searches compared to White people, which violates our predictions.

Table 23: Tukey's HSD test result (Age and Strip Search)

group1	group2	meandiff	p-adj	lower	upper	reject
17 and younger	18 - 24	0.0423	0.001	0.0226	0.0621	True
17 and younger	25 - 34	0.0402	0.001	0.0217	0.0588	True
17 and younger	35 - 44	0.0371	0.001	0.0182	0.0559	True
17 and younger	45 - 54	0.0078	0.9	-0.0122	0.0278	False
17 and younger	55 - 64	-0.0132	0.5756	-0.0355	0.0091	False
17 and younger	65 and older	-0.0648	0.001	-0.0963	-0.0334	True
18 - 24	25 - 34	-0.0021	0.9	-0.0137	0.0095	False
18 - 24	35 - 44	-0.0052	0.849	-0.0174	0.0069	False
18 - 24	45 - 54	-0.0345	0.001	-0.0484	-0.0207	True
18 - 24	55 - 64	-0.0555	0.001	-0.0725	-0.0385	True
18 - 24	65 and older	-0.1071	0.001	-0.1351	-0.0792	True
25 - 34	35 - 44	-0.0032	0.9	-0.0131	0.0068	False
25 - 34	45 - 54	-0.0324	0.001	-0.0444	-0.0204	True
25 - 34	55 - 64	-0.0534	0.001	-0.069	-0.0379	True
25 - 34	65 and older	-0.105	0.001	-0.1321	-0.078	True
35 - 44	45 - 54	-0.0293	0.001	-0.0418	-0.0168	True
35 - 44	55 - 64	-0.0503	0.001	-0.0662	-0.0343	True
35 - 44	65 and older	-0.1019	0.001	-0.1292	-0.0746	True
45 - 54	55 - 64	-0.021	0.0064	-0.0383	-0.0037	True
45 - 54	65 and older	-0.0726	0.001	-0.1007	-0.0445	True
55 - 64	65 and older	-0.0516	0.001	-0.0814	-0.0218	True

Table 23 shows the differences in strip searches across age groups. Although there were many groups whose results were statistically significantly different, most of their mean was within 0.1. Of all the statistically significant results, the differences in strip searches between people aged 65 and older and the other groups were obviously larger as all mean differences were around -0.1, meaning that older adults encountered significantly fewer strip searches compared to other age groups. Meanwhile, people aged 17 and younger experienced fewer strip searches compared to other age groups, except for those aged 55 and above. Therefore, youth and the elderly experienced statistically significantly fewer strip searches, and the elderly had the fewest.

Result

From the descriptive analysis, youth and the elderly experienced the fewest strip searches, but people of color had similar numbers of strip searches as White people. From the t-test, there were statistically significantly fewer strip searches for all minority groups (i.e., youth, seniors, and people of color). From the two-way ANOVA test, although there were many unexplained noises in the model, we found that races, age groups, and their interactions were significant predictors of strip searches because they had a significant effect on strip searches. From the two Tukey's HSD tests, people of color did not experience more strip searches than White people, but youth and the elderly had significantly fewer strip searches, which is consistent with the descriptive analysis. Hence, we can conclude that age groups may have a greater effect on strip searches than race, as evidenced by all analyses. Besides, these tests do not fit our hypothesis that the Toronto Police Service may discriminate against minority groups by conducting more strip searches of them. Instead, most of them encountered even far fewer strip searches.

Due to the inconsistent results of the various tests and the descriptive analysis, there might be other variables that influence the decision to strip search, especially the crime categories. However, it is noted that these analyses cannot show age groups and races lead to the decision of the strip search, but only that these variables are at least predictors of the strip search.

General Discussion

Results Synthesis

Overall, the project conducted study on three research question that evaluated the data from the Toronto Police Service. The first question examines whether there is a gender-based difference in strip search rates. The study finds that male individuals are more likely to be strip searched than females but does not conclude that the Toronto Police Service practices discriminatory measures. The second question looks at whether the Toronto Police Service practices discriminatory measures against youth and the elderly in making arrest decisions. The study finds that there is a significant difference in arrest rates between youth and the elderly and other middle-aged groups, but there is no evidence of a difference in arrest rates between youth and the elderly. The third question examines whether there is racial discrimination in strip search practices. The study finds that people of color had similar numbers of strip searches as White people, but youth and the elderly had significantly fewer strip searches, which suggests that age groups may have a greater

effect on strip searches than race. However, the study cannot demonstrate a causal relationship between age groups or race and strip search practices. In conclusion, the study on the Toronto Police Service's arrest and strip search practices provides insights into gender, age, and racial disparities in law enforcement. While the results suggest that male individuals are more likely to be strip searched and that youth and the elderly have different arrest rates compared to middle-aged individuals, the study cannot demonstrate a causal relationship between these factors and police practices. Therefore, further research is needed to better understand the complex factors that influence law enforcement practices and to address any potential discrimination or biases.

Limitation and Future Work

As this dataset does not contain any continuous variables, we chose the boolean attributes, including the 'Booked' and 'StripSearch', as the dependent variables for the research questions. We considered creating new columns for the dataset, such as percentage counts of female arrest rates, but it still relies on the 'Booked' column, which is more of a descriptive analysis. Since the dependent variables are not continuous, we were unable to develop different types of graphs and analyzes, such as boxplots and standard deviations, which limits the diversity of EDA. Besides, the boolean variables may reduce the accuracy of all models and tests because they can only be viewed as probabilities of occurrence rather than true continuous variables.

Additionally, there are more variables than gender, age, and race that affect arrest rates and the number of strip searches, especially the type of crime and actions at arrest. Hence, we can continue to study the relationship between more variables by building more statistical models, such as the chi-squared model of the relationship and dependence between the type of crime and race, or incorporate more variables into the two-way ANOVA model. These tests will provide a deeper understanding of their relationship.

Also, this report provides a preliminary analysis and investigation of the relationship among gender, age, race, arrest rates, and the number of strip searches. It does show that there is a strong relationship between them and predicts both dependent variables. Yet, these analyses could not draw conclusions about casual relationships. We can only know which groups have significantly higher strip searches or arrest rates, but we cannot conclude that Toronto Police Service discriminates against them.

Therefore, after confirming their relationships, we can determine causality by conducting experiments, such as randomized controlled trials, which helps to answer our research questions of finding out whether specific groups are being discriminated against by the Toronto Police Service when deciding to strip search or arrest them.

Conclusion

This report examines whether the Toronto Police Service practiced discriminatory measures against certain groups, including male individuals, youth, the elderly, and people of color, when

making strip searches and arrest decisions. Although none of the analyses were able to confirm a causal relationship, the number of strip searches were statistically significantly higher among male, while the number of strip searches and arrest rates was statistically significantly lower for youth, the elderly, and people of color. The result aligns with our initial hypothesis that these groups were statistically significantly different from the others, but we failed to predict the direction of the differences. Also, many noises affected the accuracy of our models. Hence, it is suggested that further research using more variables and different models is needed to determine causality and to see if the Toronto Police Service discriminates against any groups. We hope the Toronto Police Service can use these insights to improve its training and prevent discrimination when making decisions.

Bibliography

- [1]Adrian Ghobrial. 2021. CityNews. *toronto.citynews.ca*. Retrieved February 21, 2023 from <https://toronto.citynews.ca/2021/03/02/strip-searches-by-toronto-police-drop-dramatically-in-february/>
- [2]Wendy Gillis. 2020. “Clearly, we were doing it wrong”: Toronto police are doing far fewer strip searches under strict new rules, Interim chief says. *thestar.com*. Retrieved February 20, 2023 from <https://www.thestar.com/news/gta/2020/11/23/clearly-we-were-doing-it-wrong-toronto-police-doing-far-fewer-strip-searches-under-new-rules-interim-chief-says.html>
- [3]Monika Lemke. 2022. Policing Toronto: Strip Searching in a Divided City - The Bullet. *Socialist Project*. Retrieved February 21, 2023 from <https://socialistproject.ca/2022/07/policing-toronto-strip-searching-in-a-divided-city/>